FEBS Letters 586 (2012) 1724–1730

journal homepage: www.FEBSLetters.org

Structural and functional consequences of mutating a proteobacteria-specific surface residue in the catalytic domain of *Escherichia coli* GluRS

Saumya Dasgupta *, Debasis Manna [1], Gautam Basu *

Department of Biophysics, Bose Institute, P-1/12 CIT Scheme VIIM, Kolkata 700 054, West Bengal, India

ARTICLE INFO

Article history:
Received 17 February 2012
Revised 2 May 2012
Accepted 2 May 2012
Available online 11 May 2012

Edited by Michael Ibba

Keywords:
Glutamyl-tRNA synthetase
$tRNA^{Glu}$
Glutamylation
Proteobacteria
Phylum-specificity
Escherichia coli

ABSTRACT

Nucleotides whose mutations seriously affect glutamylation efficiency are experimentally known for *Escherichia coli* $tRNA^{Glu}$. However, not much is known about functional hotspots on the complementary enzyme, glutamyl-tRNA synthetase (GluRS). From structural and functional studies on an Arg266Leu mutant of *E. coli* GluRS, we demonstrate that Arg266 is essential for efficient glutamylation of $tRNA^{Glu}$. Consistent with this result, we found that Arg266 is a conserved signature of proteobacterial GluRS. In contrast, most non-proteobacterial GluRS contain Leu, and never Arg, at this position. Our results imply a unique strategy of glutamylation of $tRNA^{Glu}$ in proteobacteria under phylum-specific evolutionary compulsions.

1. Introduction

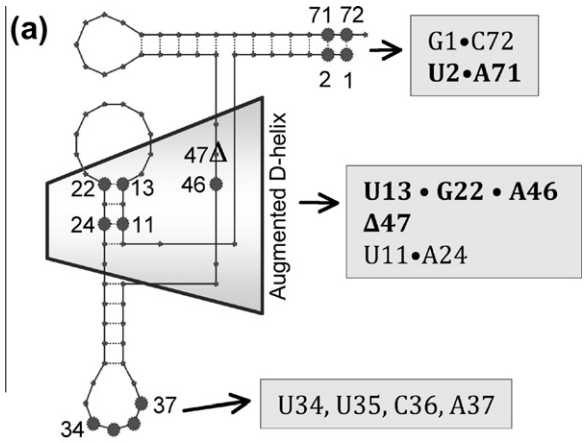
Protein translation demands high fidelity. There are number of molecular checkpoints to fulfill this demand. Among these, insuring that a particular tRNA gets aminoacylated by its conjugate amino acid, which in turn is catalyzed by a particular aminoacyl-tRNA synthetase (aaRS), is important [1]. Typically, a tRNA needs to be specific to its cognate aaRS, and vice versa. Understanding the molecular basis of specificity in the aminoacylation reaction machinery is an active field of research [2–7] that can yield important clues about molecular mechanisms of specific and non-specific aaRS–tRNA interaction.

Experimentally, one way to probe aaRS–tRNA interaction is selective mutation of nucleotides (in tRNA) or amino acid residues (in aaRS) and studying the effect of the mutations on aminoacylation efficiency. Experiments along these lines have been performed on $tRNA^{Glu}$ in the bacterium *Escherichia coli*, where $tRNA^{Glu}$ nucleotides were extensively mutated and the glutamylation efficiency of mutant $tRNA^{Glu}$, catalyzed by glutamyl-tRNA synthetase (Glu RS), were measured. This has yielded a nucleotide identity set of *E. coli* $tRNA^{Glu}$ – nucleotides that play a critical role in maintaining optimum glutamylation efficiency [2]. As shown in Fig. 1a, the identity elements in *E. coli* $tRNA^{Glu}$ are clustered in the anticodon loop, the acceptor arm and the augmented D-helix in $tRNA^{Glu}$. In this paper we focus on the identity elements present at the acceptor arm and the augmented D-helix in $tRNA^{Glu}$. Bacterial GluRS constitutes of a N-terminal catalytic domain and a C-terminal anticodon-binding domain [8]. The catalytic domain interacts with the acceptor arm and the augmented D-helix in $tRNA^{Glu}$ (Fig. 2a) [3].

Limited mutational studies have been performed on the catalytic domain in a few bacterial GluRSs. Mutation of a residue (C100Y), close to the acceptor stem of $tRNA^{Glu}$ in the Zn-binding SWIM domain in *E. coli* GluRS resulted in a variant with a slightly lower affinity for L-Glu suggesting that the SWIM domain participates in correctly positioning the tRNA acceptor end in the active site [9]. Mutational studies on *Thermus thermophilus* GluRS have also yielded a set of residues whose mutation affects the glutamylation efficiency [8]. Among the set of mutated catalytic domain residues, mutation of four residues, E282, S299, K309 and W312, were shown to play an important role in glutamylation of $tRNA^{Glu}$. All four residues are close to the D-helix of $tRNA^{Glu}$ as shown in Fig. 2a. Limited mutational studies have also been performed on GluRS1 and GluRS2 from *Helicobacter pylori* where GluRS1 stands for the canonical GluRS and GluRS2 corresponds to a non-canonical version that only glutamylates $tRNA^{Gln}$ [10]. In addition to mutational studies, glutamylation efficiencies of a domain-deleted or a domain-swapped version of *E. coli* GluRS have also been reported – from our lab [11,12] and by Lapointe and co-workers [13]. These

* Corresponding authors. Fax: +91 33 2355 3886 (G. Basu).
E-mail addresses: saumyadg@gmail.com (S. Dasgupta), gautam@boseinst.ernet.in (G. Basu).
[1] Current Address: Department of Chemistry, Indian Institute of Technology, Guwahati 781 039, Assam, India.

0014-5793/$36.00
http://dx.doi.org/10.1016/j.febslet.2012.05.006



(b)

		1–72	2–71	11–24	13–22–46	47
ECO	E1	**G–C**	**U–A**	**U–A**	**U–G–A**	**Δ**
TTH	E1	G–C	G–U	A–U	U–G–G	Δ
TTH	E2	G–C	G–U	U–A	U–G–A	Δ
TEL	E1	G–U	C–G	U–A	U–G–A	Δ
BSS	E1	G–C	G–U	U–A	A–A–A	Δ
BSS	E1	G–C	G–U	U–G	A–A–A	Δ
MTU	E1	G–U	C–G	U–A	U–G–A	Δ
MTU	E2	G–U	G–C	U–A	U–G–G	Δ

Fig. 1. *E. coli* $tRNA^{Glu}$ identity elements and their conservation. (a) A cartoon representation of $tRNA^{Glu}$ where identity nucleotides in *E. coli* $tRNA^{Glu}$ are shown as solid circles (Δ stands for absence of a nucleotide). (b) Nature of $tRNA^{Glu}$ nucleotides in *E. coli* (ECO), *T. thermophilus* (TTH), *Thermosynechococcus elongatus* (TEL), *B. subtilis* (BSS) and *M. tuberculosis* (MTU) at positions corresponding to *E. coli* $tRNA^{Glu}$ identity elements (nucleotides that are non-identical to ECO are highlighted in gray; E1 and E2 represent $tRNA^{Glu}$ (^{34}UUC) and $tRNA^{Glu}$ (^{34}CUC) respectively).

studies show that the isolated N-terminal catalytic domain of *E. coli* GluRS is still capable of glutamylating $tRNA^{Glu}$, albeit with a reduced efficiency.

Even when $tRNA^{Glu}$ identity elements are known, as in *E. coli* $tRNA^{Glu}$, there is no guarantee that the identity set is universal among all bacterial $tRNA^{Glu}$. As shown in Fig. 1b, a simple comparison of *E. coli* (gamma-proteobacteria) $tRNA^{Glu}$ identity elements with corresponding nucleotide sequences in $tRNA^{Glu}$ from four other bacteria – *T. thermophilus* (deinococcus-thermus), *T. elongatus* (cyanobacteria), *Bacillus subtilis* (firmicutes) and *Mycobacterium tuberculosis* (actinobacteria) — illustrates this point. The U2·A71 identity in the acceptor stem seems to be specific to the proteobacterium *E. coli* while a non-augmented D-helix (A13·A22) only occurs for the firmicute *B. subtilis*. It is interesting that *B. subtilis* GluRS cannot charge the augmented D-helix containing *E. coli* $tRNA^{Glu}$ [14] supporting the idea that the $tRNA^{Glu}$ identity elements in *E. coli* and *B. subtilis* are different. Indirect evidence that the distribution of tRNA identity elements in *E. coli* $tRNA^{Glu}$ may not be universal also comes from studies on GluRS2, a non-canonical GluRS, from glutamylation assays on chimeric tRNAs in *H. pylori* [15] (an epsilon proteobacteria) and from glutamylation assays on $tRNA^{Glu}$ isoacceptors of *Acidithiobacillus ferrooxidans* [16] (a gamma proteobacteria).

If some of the experimentally determined $tRNA^{Glu}$ identity elements are unique to *E. coli*, then there also must be some unique phylum-specific residues in the corresponding GluRS (*E. coli* in particular, and proteobacteria in general) as well. In this paper we focus on identifying such residues in *E. coli* GluRS. Our strategy is to utilize whole genome bacterial sequences from a large number of bacteria and compare the GluRS sequence vis-à-vis $tRNA^{Glu}$ sequences. Following this strategy, Arg266 in *E. coli* GluRS was identified to be unique in proteobacterial GluRS, changing mostly to Leu in non-proteobacteria. Arg266 is conserved in bacteria that possess both the augmented D-helix and the acceptor arm identity elements (Fig. 1a). An Arg266Leu mutant of *E. coli* GluRS was constructed to explore structural and functional role of Arg266. Replacement of Arg266 by Leu drastically altered the catalytic

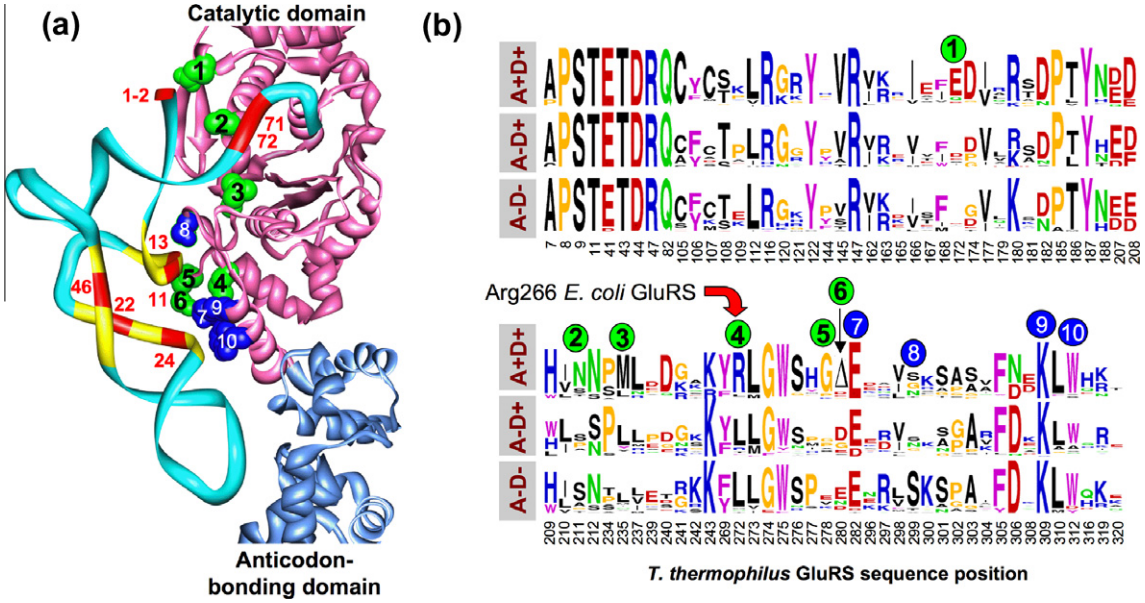


Fig. 2. Sequence conservation of $tRNA^{Glu}$-interaction residues in GluRS. (a) Cartoon representation of *T. thermophilus* GluRS–$tRNA^{Glu}$ complex (pdb code: 1n78). $tRNA^{Glu}$ is colored cyan with the augmented D-helix colored yellow and the identity elements (Fig. 1a) colored red. (b) Sequence logo plots (sequence alignment file given in Supplementary Fig. S2) of 137 bacterial GluRS sequences corresponding to 81 GluRS sequence positions that are in close proximity (within 6 Å) to $tRNA^{Glu}$ nucleotides in the crystal structure of panel (a). Grouping of GluRS sequences (A+D+, A-D+ and A-D-) are explained in the text. Residues positions in GluRS that show A+D+ group-specific conservation (green circles), and others, whose mutation seriously affected glutamylation efficiency of *T. thermophilus* GluRS (blue circles) are marked in both panels.

efficiency of GluRS in addition to slightly altering the secondary structure and the stability of the protein. The implications of our results are discussed with a focus on phylum-specific $tRNA^{Glu}$–GluRS interaction in bacteria.

2. Materials and methods

2.1. Materials

E. coli $tRNA^{Glu}$, L-Glu, ATP and ultrapure urea were purchased from Sigma. The acceptor capacity of *E. coli* $tRNA^{Glu}$ was ~1.2 nmol/OD_{260nm}. $tRNA^{Glu}$ concentration was determined assuming 1.6 nmol/ml/OD_{260nm} for 100% aminoacylation [17]. The specific radioactivity of [^{3}H] L-Glu (GE Healthcare) was 50 Ci/mmol.

2.2. Construction of R266L mutant of E. coli GluRS

Two primers, 5′CTGCTGAACTATCTGGTGCTTCTGGGCTGGTCC CACGGCGAT3′ and 5′ATCGCCGTGGGACCAGCCCAGAAGCACCAGAT AGTTCAGCAG3′, were purchased from MWG-biotech and used for *in vitro* site directed mutagenesis (Stratagene) using *E. coli* GluRS gene (plasmid PLQ7612) as the template [18]. The resultant plasmid DNA was purified by QIAGEN-plasmid purification protocol (Qiagen). The mutation (R266L) was confirmed by DNA sequencing (Applied Biosystems).

2.3. Enzyme purification

Plasmid DNAs corresponding to the wild type *E. coli* GluRS and the R266L mutant were separately transformed into *E. coli* BL21 (DE3). Single colonies from the plates were inoculated in Luria Broth (1 liter) and expressed at 37 °C at 200 rpm (with 50 μg/ml of kanamycin). At an O.D. (595 nm) of 0.5–0.6, cells were induced by 1 mM IPTG and shaken at 37 °C for six hours. Purified enzymes from the harvested pellets were obtained by Ni–NTA affinity chromatography (Qiagen) using a procedure described elsewhere [12]. The eluted protein fractions were analyzed by 12% SDS–PAGE. After dialysis in 50 mM phosphate buffer (pH 7.5 containing 100 mM KCl, 10 mM β-mercaptoethanol and 10% glycerol) the purified protein fractions were stored at −80 °C.

2.4. Binding studies

Binding of $tRNA^{Glu}$ and ATP with wild-type *E. coli* GluRS and the R266L mutant was monitored by Trp-fluorescence quenching at 24 °C in a Hitachi F7000 spectrofluorimeter in 20 mM HEPES buffer (pH 7.5) containing 5 mM $MgCl_2$ and 10% glycerol using a protocol described before [12]. Protein concentrations were 0.5 μM for $tRNA^{Glu}$ titrations and 2 μM for ATP titrations. The excitation and emission wavelengths were 295 and 340 nm respectively with appropriate correction for inner filter effect. The resulting binding isotherms were analyzed assuming a 1:1 binding stoichiometry [19].

2.5. Glutamylation assay

In vitro glutamylation assay of the enzymes were carried out with 5 μM *E. coli* $tRNA^{Glu}$ in 50 mM HEPES (pH 7.5), 0.1 mM unlabeled L-Glu, 16 mM $MgCl_2$, 2 mM ATP, 0.8 mM β-mercaptoethanol and [^{3}H] L-Glu (1 μl of stock per 100 μl assay buffer) at 37 °C as described before [12]. Protein concentrations used per assay point were 1.8 nM (*E. coli* GluRS) and 0.36 μM (the R266L mutant). Kinetic parameters (K_m and k_{cat}) associated with the glutamylation reactions were determined with respect to $tRNA^{Glu}$ (0–8 μM) at 37 °C as described earlier [12].

2.6. Structural studies by CD and fluorescence spectroscopy

Far UV circular dichroism (CD) studies were performed on a Jasco J-815 spectro-polarimeter at 25 °C in 50 mM phosphate buffer (pH 7.5) containing 100 mM KCl and 10% glycerol. Steady state fluorescence spectra were obtained in the same buffer in a LS55 Fluorescence Spectrometer (Perkin Elmer). Cuvette path lengths were 2.0 and 5.0 mm for CD and fluorescence studies, respectively.

2.7. Equilibrium unfolding studies

Urea stock solutions were prepared in 50 mM potassium phosphate buffer (pH 7.5), containing 100 mM KCl and 10% glycerol. All protein samples in urea (final concentration 2 μM) were incubated for ~16 h before spectroscopic measurements at 24 °C. The ratio of fluorescence intensities (F_{340}/F_{350}; at 340 and 350 nm) and ellipticity at 222 nm (Θ_{222}) were plotted against urea concentrations and were fitted to a two-state model. Thermal denaturation was monitored by CD spectroscopy (Θ_{222}) in a Jasco J-815 spectropolarimeter attached with a temperature-controlled peltier.

2.8. Sequence analysis of GluRS and $tRNA^{Glu}$

A database comprising bacterial GluRS and $tRNA^{Glu}$ sequences was constructed from 137 bacterial species covering a wide range of bacterial phyla avoiding genus duplication and excluding bacteria with two copies of GluRS. The sequences were downloaded from KEGG genomic database [20] and are listed in Supplementary Tables S1–S12 and Supplementary Fig. S1. GluRS sequences were aligned using MUSCLE [21] and sequence conservation viewed using JALVIEW [22]. $tRNA^{Glu}$ sequences were aligned manually (Supplementary Fig. S1). Aligned sequences were visualized as sequence logos using WEBLOGO [23]. Structures were visualized and analyzed using the program Chimera [24] and structural alignments were performed using the program MATRAS [25]. Phylogenetic analysis of GluRS sequences was performed using the online server phylogeny.fr using the "one-click" mode [26].

3. Results and discussion

3.1. Arg266 of E. coli GluRS correlates with E. coli $tRNA^{Glu}$ identity elements

The experimentally determined identity elements in *E. coli* $tRNA^{Glu}$ [2] are distributed in three distinct regions – the acceptor stem (G1·C72, U2·A71), the anticodon loop (U34, U35, C36, A37) and the augmented D-helix (U11·A24, U13·G22··A46, and Δ47), as shown in Fig. 1a. Our aim is to identify unique amino acids in *E. coli* GluRS that correlate with *E. coli* $tRNA^{Glu}$ identity elements. Based on the presence of two *E. coli* $tRNA^{Glu}$ identity features in the $tRNA^{Glu1}$ (^{34}UUC^{36}), the A2·U71 (or U2·A71) base-pair in the acceptor arm (signature 'A') and the augmented D-helix (U13·G22··A46 + Δ47) (signature 'D'), 137 GluRS sequences in our database were grouped into three classes. These are: (i) A+D+ group (bacteria possessing both A and D signatures), (ii) A-D+ group (bacteria lacking signature A but possessing signature D) and, (iii) A-D- group (bacteria lacking both the signatures). Crystal structure of *T. thermophilus* $tRNA^{Glu}$–GluRS complex [27] show that the N-terminal catalytic domain of GluRS is in proximity to the two *E. coli* $tRNA^{Glu}$ identity features. In the absence of free or $tRNA^{Glu}$-bound structure of *E. coli* GluRS, the $tRNA^{Glu}$-bound structure of *T. thermophilus* GluRS [3] was examined and catalytic

domain (1–322) residues [8] in *T. thermophilus* GluRS that show close contact with $tRNA^{Glu}$ (within 6 Å) were identified. A total of 81 GluRS residues were identified based on this criterion. A complete list of the residues along with interacting $tRNA^{Glu}$ nucleotides is given in Supplementary Table S13.

Amino acids that occupy the 81 $tRNA^{Glu}$-interacting sequence positions in *T. thermophilus* GluRS were extracted from multiple sequence alignment of 137 bacterial GluRS sequences and grouped according to scheme described above (A+D+, A-D+ and A-D-). The corresponding sequence logo plot is shown in Fig. 2b (the complete sequence alignment is given in Supplementary Fig. S2). We sought to identify amino acid residues in the A+D+ group (corresponding to *E. coli*) whose mutation may make *E. coli* GluRS incompatible with the A+D+ signature of *E. coli* $tRNA^{Glu}$. As shown in Fig. 2b, the A+D+ group exhibited unique residue conservation (>60%) at a number of sequence positions (E172, N211, M235, R272, G278, Δ280; also see Supplementary Fig. S3). Of these, only two positions, 235 and 272, also showed residue conservation in the A-D+ and the A-D- groups (L235 and L272, respectively) that was different from the A+D+ group (see Supplementary Fig. S3). The nature of the conserved amino acids in the A+D+ group is very different from that in the A-D+/A-D- group at position 272 (Arg/Leu) compared to position 235 (Met/Leu). Arginine, present at 272 in the A+D+ group, is also known to play an important role in protein–nucleic acid interaction [28]. This led us to focus on position 272 in *T. thermophilus* GluRS. GluRS from *E. coli*, a gamma-proteobacteria, contains Arg (Arg266) at the equivalent position 266 (unless otherwise stated, GluRS residue numbering in the rest of the manuscript will correspond to *E. coli* GluRS sequence numbering). An R266L mutant of *E. coli* GluRS was constructed to study the structural and functional effects of the mutation.

3.2. R266L mutant exhibits impaired $tRNA^{Glu}$- and ATP-binding

In order to glutamylate $tRNA^{Glu}$, GluRS must form a binary complex with $tRNA^{Glu}$. Binding of $tRNA^{Glu}$ is known to quench the intrinsic fluorescence of GluRS [12]. We estimated the dissociation constants (K_d) associated with $tRNA^{Glu}$ binding to the wild type and the R266L mutant of *E. coli* GluRS by titrating $tRNA^{Glu}$ against the proteins and measuring the intrinsic Trp fluorescence quenching (Fig. 3a). The measured K_d values for wild type GluRS and the R266L mutant were 59 ± 5 nM and 306 ± 3 nM, respectively, demonstrating that R266L mutation induces a 5-fold decrease in the $tRNA^{Glu}$-affinity of *E. coli* GluRS. Binding of ATP, another substrate for GluRS, also reduces the intrinsic fluorescence of GluRS [12]. ATP binding to GluRS and the R266L mutant, monitored by quenching of Trp fluorescence (Fig. 3b), yielded a larger dissociation constant (K_d = 112 ± 6 μM) for the R266L mutant than GluRS (K_d = 31 ± 3 μM), indicating that R266L mutation reduces GluRS affinity for ATP by about 3-fold. In summary, the R266L mutation impairs binding of two essential substrates to GluRS, $tRNA^{Glu}$ and ATP. It was also observed that the extent of fluorescence quenching brought about by both ATP and $tRNA^{Glu}$, at saturating concentrations, were much larger for the R266L mutant when compared to the wild type GluRS. This indicates that ATP and $tRNA^{Glu}$ binding perturbs the environment of Trp residues differently in wild type GluRS and the R266L mutant.

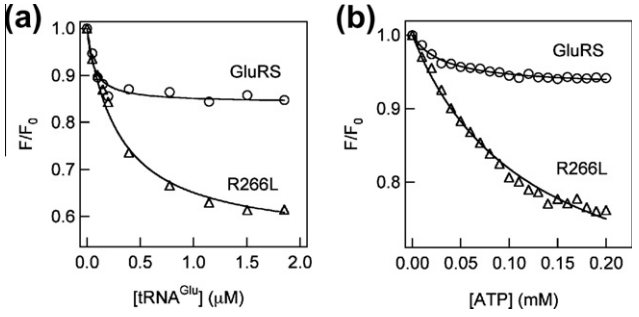


Fig. 3. $tRNA^{Glu}$ and ATP binding of *E. coli* GluRS and the R266L mutant. (a) Intrinsic Trp fluorescence quenching of *E. coli* GluRS and R266L as a function of: (a) $tRNA^{Glu}$ and (b) ATP.

3.3. R266L mutant weakened $tRNA^{Glu}$-glutamylation efficiency

Kinetic studies were performed to probe $tRNA^{Glu}$-glutamylation efficiencies of GluRS and the R266L mutant. The experimental K_m (with respect to $tRNA^{Glu}$) and k_{cat} values for GluRS were 0.3 ± 0.1 μM and 6.4 ± 1.1 s^{-1}, respectively. The R266L mutant was found to glutamylate $tRNA^{Glu}$ with a lower efficiency than wild type GluRS, reflected both in K_m (9.9 ± 0.5 μM, a 33-fold increase) and in k_{cat} ($8.2 \times 10^{-2} \pm 0.4 \times 10^{-2}$ s^{-1}, a 78-fold decrease). Taken together, in terms of k_{cat}/K_m ratio, the R266L mutant is about 2570-fold less efficient than wild type GluRS. Arg266 (Leu272 in *T. thermophilus* GluRS structure) is present at the $tRNA^{Glu}$–GluRS interface. Therefore, the increase in K_d or K_m is understandable. However, because Leu272 is at least 20 Å away from the bound glutamate in the crystal structure (pdb code: 1N78), the effect on the catalytic step (k_{cat}) upon mutating this residue is counterintuitive. Clearly long-range communication between this site and the catalytic site must be operative. In addition, since R266L mutation affected ATP binding, it is quite likely that the altered ATP binding also plays a role in lowering k_{cat}.

3.4. The R266L mutant is more stable than wild type GluRS and exhibits reduced helical content

Justifying the observed alteration in glutamylation efficiency of the R266L mutant using a structure based model assumes that *E. coli* GluRS structure is identical to that of *T. thermophilus* GluRS and that the structure is unaffected by the R266L mutation. To confirm this, especially the second assumption, comparative fluorescence and CD spectroscopic studies were performed on *E. coli*

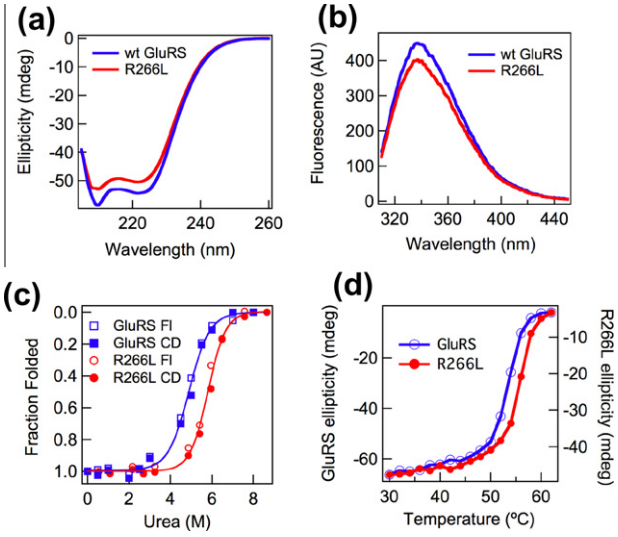


Fig. 4. Spectroscopic signatures and unfolding profiles of GluRS and the R266L mutant. CD spectra (a), fluorescence spectra (b), urea-induced unfolding profiles (c), monitored by CD (Θ_{222}) and fluorescence (F_{340}/F_{350}), and, temperature-induced unfolding profiles (d), monitored by CD, of *E. coli* GluRS and the R266L mutant. The solid lines in panel (c) correspond to two state fits (ΔG° and m for unfolding were 5.8 ± 0.5 kcal/mol and 1.2 ± 0.1 kcal/mol/M, respectively, for wt GluRS, and, 8.0 ± 0.5 kcal/mol and 1.4 ± 0.1 kcal/mol/M, respectively, for R266L).

GluRS and the R266L mutant. In Fig. 4a, the far-UV CD spectra of *E. coli* GluRS along with the R266L mutant are shown. Compared to *E. coli* GluRS, the R266L mutant exhibits a slightly diminished negative ellipticity (222 nm) that translates into a slight loss of helicity. The fluorescence spectra for *E. coli* GluRS and the R266L mutant are shown in Fig. 4b. The emission maxima of the two proteins are identical (~332 nm), indicating that the overall solvent exposures of Trp residues are unaffected by the R266L mutation. However, the R266L mutant exhibited a slightly decreased fluorescence intensity indicating that the R266L mutation may affect the non-radiative de-excitation pathways of nearby Trp residues, which could be a consequence of the CD-detected altered secondary structure of R266L.

Urea-induced equilibrium unfolding profiles (Fig. 4c) showed that the R266L mutant unfolds at a slightly higher urea concentration than *E. coli* GluRS, the R266L mutant being about 2 kcal/mol more stable (see legend to Fig. 4c for thermodynamic parameters). The extra stability of the R266L mutant was also observed in temperature-induced equilibrium unfolding studies. As can be seen from Fig. 4d, the T_m value for the R266L mutant is about 3 degrees higher than the T_m for *E. coli* GluRS. Unfortunately, detailed thermodynamic analyses of the temperature-induced unfolding profiles could not be pursued since both the proteins precipitated beyond 60 °C.

3.5. The loop following R266 in E. coli GluRS is uniquely different than the loop in L266 containing GluRS

Since L272 in *T. thermophilus* GluRS, the residue equivalent to R266 in *E. coli* GluRS, is a surface residue in the crystal structure (pdb code: 1n78), R266L mutation most likely diminishes the helical content of *E. coli* GluRS locally. GluRS sequences of all available crystal structures, from *T. thermophilus*, *T. elongatus*, *T. maritima* and *M. tuberculosis* (pdb codes 1n78, 2cfo, 2o5r and 2ja2, respectively) contain Leu at position 266, present at the C-terminal end of a helix (H9), followed by a loop and a helix (H10) (Fig. 5a). Structural alignment of the four structures showed that the backbone structure is conserved in this region (Fig. 5b). However, sequence analysis for the H9-loop-H10 structural segment, in 137 bacterial GluRS (Supplementary Fig. S4), showed that the loop following position 266 is uniquely different for GluRS with Arg266 as opposed to GluRS without Arg266 (mostly Leu266). The results are summarized in a sequence logo plot (Fig. 5c). The presence of Arg266 is accompanied by a four-residue signature sequence (group RΔ, Fig. 5c), ^{271}HGDΔ where Δ signifies deletion. The deletion associated with the RΔ group is absent in GluRS sequences that lack R266 (group LX, Fig. 5c), although they do not show a conserved sequence pattern in this region. The conserved GD sequence motif present in the loop in the RΔ group of GluRS, a motif associated with high turn propensity [29], and a conserved deletion following the GD sequence motif, the loop structure in Arg266-containing GluRS is likely to be different than the loop in the four GluRS crystal structures, all containing Leu266. Arg266 also must play a crucial role in stabilizing the altered loop structure in *E. coli* GluRS since its mutation to Leu slightly reduced GluRS helical content and brought about an extra stability to the protein. Four GluRS sequences (all gamma-proteobacterial) do not follow the above pattern (group RE in Fig. 5c and Supplementary Fig. S4). These GluRS sequences show characteristics of both the RΔ (R266) and the LX (lacking the deletion characteristic of the RΔ group) groups with the signature ^{271}MPDE. As shown later, these proteobacterial GluRS sequences neighbor non-proteobacterial GluRS (chlamydiae) in the phylogenetic tree of GluRS (Fig. 6) and their loop structure might be uniquely different than others.

Without a crystal structure, one can only speculate the nature of the altered loop structure in *E. coli* GluRS and how Arg266Leu mutation can disrupt the structure of the loop. In tRNA^{Glu}-bound crystal structure of *T. thermophilus* GluRS (Fig. 5d), E282 (E275 in *E. coli* GluRS and conserved in all GluRS) forms a H-bond with A24 from tRNA^{Glu}. The side-chain from E282 is spatially proximal to L272 (R266 in *E. coli* GluRS). Therefore, the E275-R266 side-chain interaction may play a role in altering the loop structure in *E. coli* GluRS. The ^{271}HGDQE^{275} sequence motif in *E. coli* GluRS (group RΔ), with a tight turn signature at GD [29], can also bring the side-chains of His271 and Glu275 in proximity forming a full turn of helix, whose stability may be modulated by Arg266. The Arg266Leu mutation possibly disrupts this helix, explaining the CD data. As shown in Fig. 5d, R297 (A290 in *E. coli* GluRS) in *T. thermophilus* GluRS forms two H-bonds with tRNA^{Glu} (nucleotides A14 and G15) and is tethered by the loop (S276). If the loop structure in *E. coli* GluRS is different than what is observed from *T. thermophilus* GluRS, it would have a direct effect on this GluRS–tRNA^{Glu} interaction. Indeed, position 290 is populated mostly by R/K/D for Leu266-containing GluRS and conserved in a phylum-specific manner (Supplementary Fig. S4a). On the other hand, for Arg266-containing GluRS, this position is not conserved (Supplementary Fig. S4b).

3.6. GluRS–tRNA^{Glu} interaction is different in proteo- and non-proteobacteria

Leu272 (Arg266 in *E. coli*) in *T. thermophilus* GluRS–tRNA^{Glu} structure (pdb ID: 1n78) is in close proximity to the augmented D-helix of RNA^{Glu}, containing major identity determinants [2]. Since R266L mutation in *E. coli* GluRS seriously affected cognate glutamylation efficiency, one is tempted to conclude that the mutation specifically disrupts interactions between R266 and the augmented D-helix of tRNA^{Glu}. If true, then the presence of Arg266 in *E. coli* GluRS is to optimize this interaction. However, analysis of bacterial GluRS sequences is inconsistent with this conclusion since it was found that many non-proteobacteria with Leu266-containing GluRS also possess tRNA^{Glu} with an augmented D-helix (Supplementary Fig. S1). Therefore, it is fair to conclude that both Arg266 and Leu266 are compatible with an augmented D-helix of tRNA^{Glu}, each associated with unique loop architectures.

Arginine never occurs at position 266 in non-proteobacterial GluRS. On the other hand, except for a few alpha- and delta-proteobacteria (shown at the bottom of Supplementary Fig. S4a), all proteobacterial GluRS sequences in our database possess Arg at position 266. This indicates that proteo- and non-proteobacterial GluRSs may use different strategies to optimize their interaction with tRNA^{Glu}. The few alpha- and delta-proteobacterial GluRS sequences that do not possess Arg266 actually cluster with non-proteobacterial GluRS sequences in a phylogenetic tree of bacterial GluRS (Fig. 6) and are present in the LX group (Supplementary Fig. S4a), indicating a different evolutionary history for these GluRSs compared to other proteobacterial GluRS, possibly involving horizontal gene transfer [30,31]. In fact, the four gamma-proteobacterial Arg266-containing GluRS sequences, that did not exhibit the characteristic ^{271}HGDΔ sequence motif (group RE, Supplementary Fig. S4c), also cluster with non-proteobacterial GluRS in the phylogenetic tree, suggesting horizontal gene transfer [32,33]. However, unlike the alpha- and delta-proteobacterial GluRS, these gamma-proteobacterial GluRS have acquired Arg266, with a unique loop sequence motif ^{271}MPDE.

A detailed discussion of the evolutionary compulsions that may have driven, not all but most, proteobacterial GluRS to acquire Arg at position 266 is beyond the scope of this article although preliminary studies (Dasgupta & Basu, unpublished data) indicate that Arg266 was incorporated in proteobacterial GluRS for tRNA^{Gln}-discrimination. In absence of a more rigorous sequence analysis, here we propose that Arg266 was acquired by most proteobacterial GluRS, as it evolved from non-proteobacterial GluRS, under some

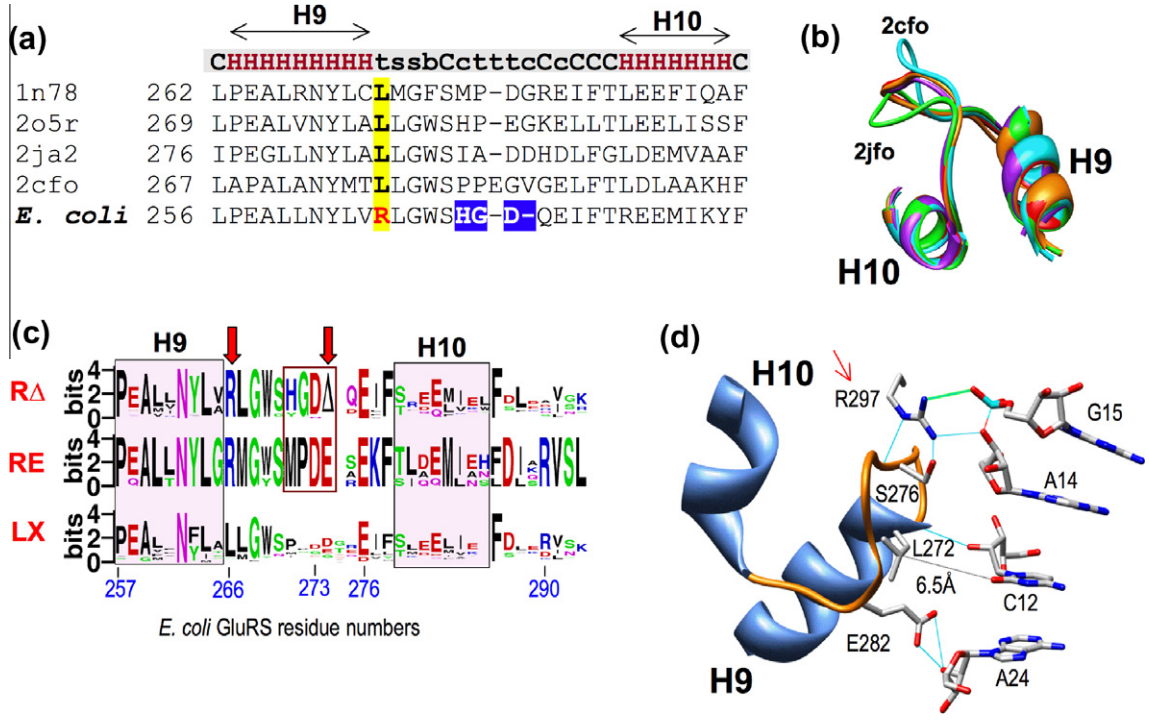


Fig. 5. Sequence and structure of the loop connecting helix 9 and helix 10 in GluRS. Sequence (a) and structure (b) alignment of H9-loop-H10 segment of four bacterial GluRS (pdb codes: 1n78, 2o5r, 2ja2 and 2cfo). (c) Sequence logo plots (sequence alignment file given in Supplementary Fig. S4) of H9-loop-H10 segment in 137 bacterial GluRS sequences (Δ indicates gap). Sequences are grouped as: RΔ (GluRS with R266 and a gap following residue 273), RE (GluRS with R266 and without a gap following residue 273), LX (GluRS without R266 and without a gap following residue 273). (d) Cartoon diagram of H9-loop-H10 region of *T. thermophilus* GluRS (pdb code: 1n78) along with H-bond interactions with $tRNA^{Glu}$. Residue R297 (not part of H9-loop-H10) is marked with a red arrow. Residue numbers correspond to *T. thermophilus* GluRS.

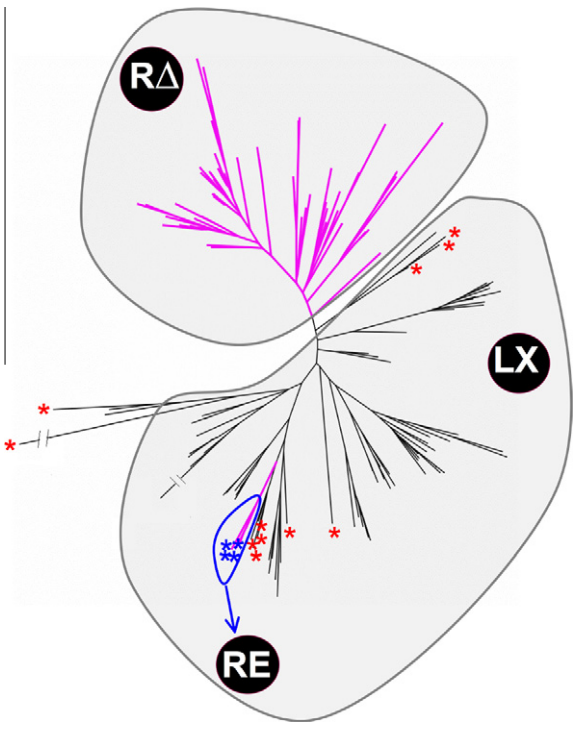


Fig. 6. Phylogenetic tree of 137 bacterial GluRS sequences. Magenta colored backbone or branches terminate in Arg266-containing GluRS. Details of the grouping scheme (RE, LX and RΔ) are described in legend to Fig. 5. Red asterisk indicates alpha- and delta-proteobacterial GluRS in the LX group and blue asterisk indicates gamma-proteobacterial GluRS in the RE group. For a more detailed annotation see Supplementary Fig. S5.

unique evolutionary pressure, not related to its interaction with $tRNA^{Glu}$. However, to compensate for any detrimental effect that this single residue change might have had on GluRS–$tRNA^{Glu}$ interaction, the loop following Arg266, and possibly some other residues, also underwent correlated change.

In this work we set out to identify *E. coli* (proteobacteria) specific residues in the catalytic domain in GluRS that correlate with the acceptor stem and the augmented D-helix identity elements of *E. coli* $tRNA^{Glu}$. As shown in Fig. 2, amino acids at a number of sequence positions exhibited unique conservation pattern in proteobacterial GluRS. Among these E172, M235 and N211 are close to the $tRNA^{Glu}$ acceptor arm while R272, G278 and Δ280 are close to the $tRNA^{Glu}$ augmented D-helix (*T. thermophilus* GluRS sequence numbering). Among these only R266 in *E. coli* GluRS (L272 in *T. thermophilus* GluRS) was mutated to L266 in the present study. However, since in the A+D+ group Arg appears at this position in a correlated fashion with the other five sequence signatures (Fig. 5b), it is quite likely that the other five residues also play an important role in efficient glutamylation of proteobacterial GluRS. Future experimental work can reveal their structural and functional roles. Also, it should be pointed out that this work was restricted to only $tRNA^{Glu}$-contacting GluRS residues, implicitly assuming that GluRS residues that do not show crystal contact may not be important. This may not be strictly correct.

Another interesting issue is the consequence of reverse mutation, namely the role of L266R mutation in non-proteobacterial GluRSs. Mutational studies have shown that four positions in *T. thermophilus* GluRS, a bacterium from the deinococcus-thermus phylum and belonging to the A-D+ group (see Fig. 2), play a pivotal role in glutamylation efficiency [8]. These residues were chosen based on their proximity to $tRNA^{Glu}$ and high conservation across bacterial GluRS sequences, rather than their selective conservation within the phylum. Our work suggests that experimental study of

L272R or G280Δ mutants (sequence positions corresponding to *T. thermophilus* GluRS) of *T. thermophilus* GluRS can provide important clues about phylum-specificity (L272 and G280 show deinococcus-thermus phylum specific conservation; see Supplementary Fig. S4a) of $tRNA^{Glu}$ glutamylation, especially since *T. thermophilus* GluRS is capable of glutamylating both *T. thermophilus* and *E. coli* $tRNA^{Glu}$ [34].

4. Conclusion

We have identified a conserved catalytic domain residue in proteobacterial GluRS, present at the $tRNA^{Glu}$–GluRS interface. The conserved residue (Arg266 in *E. coli* GluRS) is a proteobacterial signature and is correlated with the presence of a select set of *E. coli* $tRNA^{Glu}$ identity elements (augmented D-helix and U2·A71 in the acceptor stem). In non-proteobacteria the residue position is mostly occupied by Leu (but never by Arg). An R266L mutant of *E. coli* GluRS exhibited impaired glutamylation efficiency, slightly reduced helical secondary structure and enhanced stability, establishing that unlike non-proteobacterial GluRS, Leu is structurally and functionally (possibly due to local structural changes) incompatible with *E. coli* GluRS. Sequence analysis of a large database of bacterial GluRS sequences showed that Arg266 is always accompanied by a unique loop sequence that follows it, when compared to the loop that follows Leu266. Since D-helix containing $tRNA^{Glu}$ is not only present in bacteria that possess GluRS with Arg266, but also in bacteria that possess GluRS with Leu266, we propose that the true role of Arg266 is not to optimize interactions of GluRS with an augmented D-helix of $tRNA^{Glu}$. Some other unique proteobacteria-specific evolutionary pressure led to the appearance of Arg266 in proteobacterial GluRS.

Acknowledgments

This work was supported by Grants from CSIR, India. G.B. acknowledges Siddhartha Roy for critical comments and discussions. The plasmid PLQ7612 was obtained from Siddhartha Roy, originally a gift from Jacques Lapointe.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.febslet.2012.05.006.

References

[1] Ling, J., Reynolds, N. and Ibba, M. (2009) Aminoacyl-tRNA synthesis and translational quality control. Annu. Rev. Microbiol. 63, 61–78.
[2] Sekine, S. et al. (1996) Major identity determinants in the "augmented D helix" of $tRNA^{Glu}$ from *Escherichia coli*. J. Mol. Biol. 256, 685–700.
[3] Sekine, S., Nureki, O., Shimada, A., Vassylyev, D.G. and Yokoyama, S. (2001) Structural basis for anticodon recognition by discriminating glutamyl-tRNA synthetase. Nat. Struct. Biol. 8, 203–206.
[4] Ibba, M., Hong, K.W., Sherman, J.M., Sever, S. and Soll, D. (1996) Interactions between tRNA identity nucleotides and their recognition sites in glutaminyl-tRNA synthetase determine the cognate amino acid affinity of the enzyme. Proc. Natl. Acad. Sci. USA 93, 6953–6958.
[5] Perona, J.J. and Hou, Y.M. (2007) Indirect readout of tRNA for aminoacylation. Biochemistry 46, 10419–10432.
[6] Beebe, K., Mock, M., Merriman, E. and Schimmel, P. (2008) Distinct domains of tRNA synthetase recognize the same base pair. Nature 451, 90–93.
[7] O'Donoghue, P., Sheppard, K., Nureki, O. and Soll, D. (2012) Rational design of an evolutionary precursor of glutaminyl-tRNA synthetase. Proc. Natl. Acad. Sci. USA 108, 20485–20490.
[8] Nureki, O. et al. (1995) Architectures of class-defining and specific domains of glutamyl-tRNA synthetase. Science 267, 1958–1965.
[9] Banerjee, R., Dubois, D.Y., Gauthier, J., Lin, S.X., Roy, S. and Lapointe, J. (2004) The zinc-binding site of a class I aminoacyl-tRNA synthetase is a SWIM domain that modulates amino acid binding via the tRNA acceptor arm. Eur. J. Biochem. 271, 724–733.
[10] Lee, J. and Hendrickson, T.L. (2004) Divergent anticodon recognition in contrasting glutamyl-tRNA synthetases. J. Mol. Biol. 344, 1167–1174.
[11] Saha, R., Dasgupta, S., Basu, G. and Roy, S. (2009) A chimaeric glutamyl:glutaminyl-tRNA synthetase: implications for evolution. Biochem. J. 417, 449–455.
[12] Dasgupta, S., Saha, R., Dey, C., Banerjee, R., Roy, S. and Basu, G. (2009) The role of the catalytic domain of *E. coli* GluRS in $tRNA^{Gln}$ discrimination. FEBS Lett. 583, 2114–2120.
[13] Dubois, D.Y., Blais, S.P., Huot, J.L. and Lapointe, J. (2009) A C-truncated glutamyl-tRNA synthetase specific for $tRNA^{Glu}$ is stimulated by its free complementary distal domain: mechanistic and evolutionary implications. Biochemistry 48, 6012–6021.
[14] Lapointe, J., Duplain, L. and Proulx, M. (1986) A single glutamyl-tRNA synthetase aminoacylates $tRNA^{Glu}$ and $tRNA^{Gln}$ in *Bacillus subtilis* and efficiently misacylates *Escherichia coli* $tRNA^{Gln}1$ in vitro. J. Bacteriol. 165, 88–93.
[15] Chang, K.M. and Hendrickson, T.L. (2009) Recognition of $tRNA^{Gln}$ by Helicobacter pylori GluRS2–a $tRNA^{Gln}$-specific glutamyl-tRNA synthetase. Nucleic Acids Res. 37, 6942–6949.
[16] Salazar, J.C., Ahel, I., Orellana, O., Tumbula-Hansen, D., Krieger, R., Daniels, L. and Soll, D. (2003) Coevolution of an aminoacyl-tRNA synthetase with its tRNA substrates. Proc. Natl. Acad. Sci. USA 100, 13863–13868.
[17] Nissan, T.A., Oliphant, B. and Perona, J.J. (1999) An engineered class I transfer RNA with a class II tertiary fold. RNA 5, 434–445.
[18] Liu, J. et al. (1995) The zinc-binding site of *Escherichia coli* glutamyl-tRNA synthetase is located in the acceptor-binding domain. Studies by extended X-ray absorption fine structure, molecular modeling, and site-directed mutagenesis. J. Biol. Chem. 270, 15162–15169.
[19] Roy, S. (2004) Fluorescence quenching methods to study protein-nucleic acid interactions. Methods Enzymol. 379, 175–187.
[20] Kanehisa, M. and Goto, S. (2000) KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 28, 27–30.
[21] Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.
[22] Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M. and Barton, G.J. (2009) Jalview Version 2–a multiple sequence alignment editor and analysis workbench. Bioinformatics 25, 1189–1191.
[23] Crooks, G.E., Hon, G., Chandonia, J.M. and Brenner, S.E. (2004) WebLogo: a sequence logo generator. Genome Res. 14, 1188–1190.
[24] Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612.
[25] Kawabata, T. (2003) MATRAS: A program for protein 3D structure comparison. Nucleic Acids Res. 31, 3367–3369.
[26] Dereeper, A. et al. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. 36, W465–W469.
[27] Sekine, S., Nureki, O., Dubois, D.Y., Bernier, S., Chenevert, R., Lapointe, J., Vassylyev, D.G. and Yokoyama, S. (2003) ATP binding by glutamyl-tRNA synthetase is switched to the productive mode by tRNA binding. EMBO J. 22, 676–688.
[28] Jones, S., Daley, D.T., Luscombe, N.M., Berman, H.M. and Thornton, J.M. (2001) Protein-RNA interactions: a structural analysis. Nucleic Acids Res. 29, 943–954.
[29] Hutchinson, E.G. and Thornton, J.M. (1994) A revised set of potentials for beta-turn formation in proteins. Protein Sci. 3, 2207–2216.
[30] Wagner, A. and de la Chaux, N. (2008) Distant horizontal gene transfer is rare for multiple families of prokaryotic insertion sequences. Mol. Genet. Genomics 280, 397–408.
[31] Huntley, S. et al. (2011) Comparative genomic analysis of fruiting body formation in Myxococcales. Mol. Biol. Evol. 28, 1083–1097.
[32] Jeong, H. et al. (2005) Genomic blueprint of Hahella chejuensis, a marine microbe producing an algicidal agent. Nucleic Acids Res. 33, 7066–7073.
[33] Qiu, X., Kulasekara, B.R. and Lory, S. (2009) Role of horizontal gene transfer in the evolution of *Pseudomonas aeruginosa* virulence. Genome Dyn. 6, 126–139.
[34] Miyazawa, T. and Yokoyama, S. (1985) Conformational aspects and functions of tRNA. Proc. Int. Symp. Biomol. Struct. Interactions, Suppl. J. Biosci. 8, 731–737.